Improvement of Processes for Flow Information

NTT Network Service System Laboratories, NTT Corporation

Hitoshi Irino,
Masaru Katayama

NTT Network System Laboratories

Abstract of this presentation

- Ideas for increasing (optimizing) performances of processes in IPFIX
- Ideas based on all processes using **an order rule of Information Elements/fields**
- These ideas are introduced:
  - Method for **reducing the number of comparisons** between an existing flow and an incoming new packet **in Metering Processes** (MPs) (**Comparison method for multiple fields in MPs**)
  - Method for **reducing the number of copies** of flow records from Metering Process to **Exporting Processes** (EPs) with a predefined order of fields (**Copy method for multiple fields in EPs**)
  - Method for **increasing processing speed for storing** data in incoming packets to file with a predefined format of **Collecting Processes** (CPs) (**Copy method for multiple fields in CPs**)

These are basically the same.

Motivation of this research

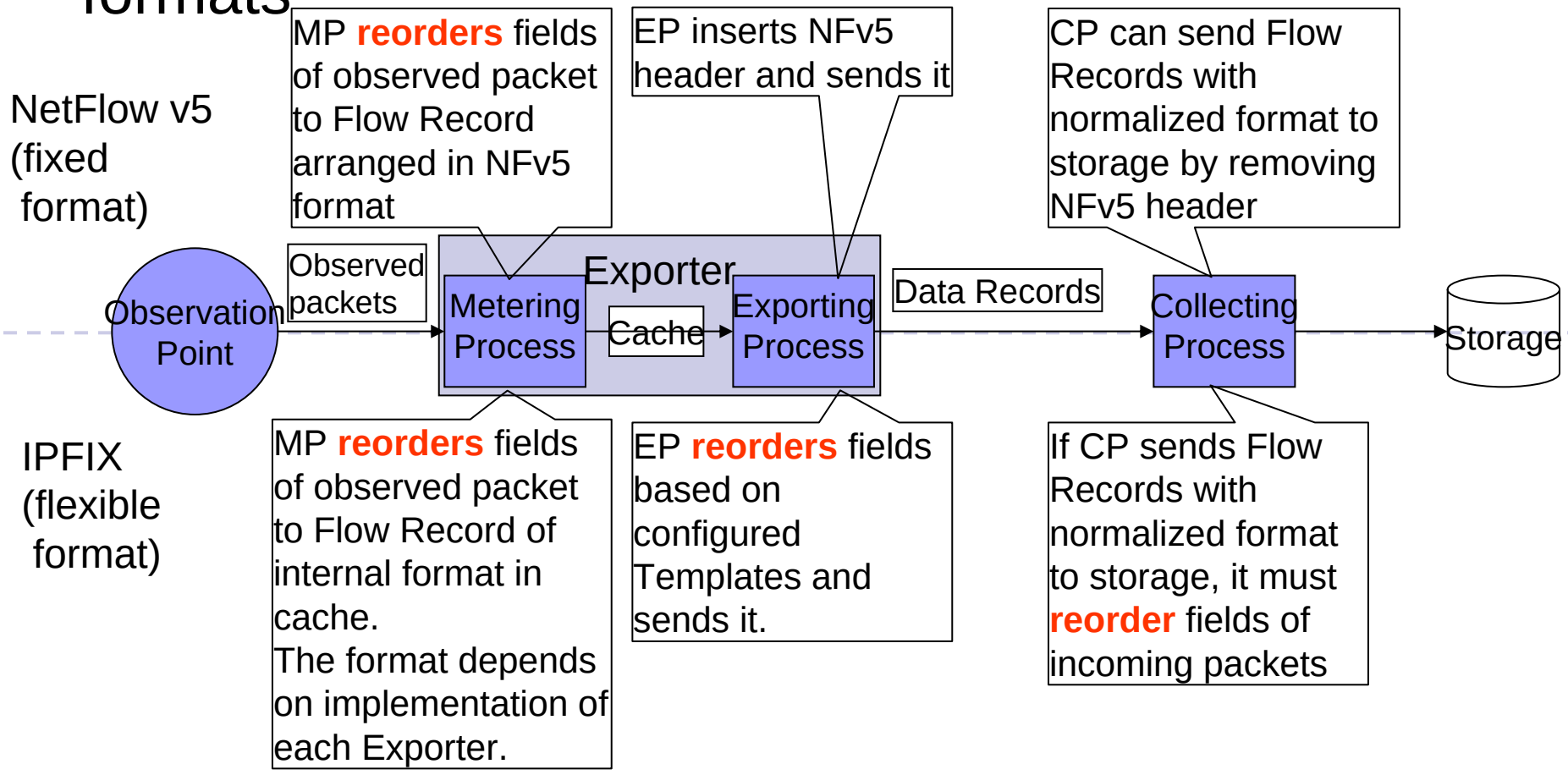
- Background
  - Network bandwidth will continue to increase.
  - IPFIX will be a standard protocol for flow information exchange.
    - Network bandwidth will become broader-band.
      - Use a lower sampling rate.
      - Use fewer Flow Keys.

→ However, flow information will become less accurate.

→ Research on increasing (optimizing) the performances of IPFIX processes

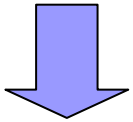
IPFIX features

- IPFIX
  - Advantage: Uses Template-based flexible flow export
  - Disadvantage: More complex than fixed-format protocol
- Comparison of processes between flexible and fixed formats

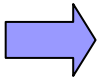


Our approach: **Making the order rule for Information Elements**

- Processes of IPFIX have a high possibility of reordering fields.
  - Reducing the cost of reordering fields can improve their performance.



- Our approach
  - Make the order rule for Information Elements
    - Order rule gives IPFIX processes chances to process multiple fields.
    - Processing multiple fields at a time achieves higher performance than processing one field at a time.
    - The rule does not influence the flexibility of IPFIX.



If a unified order rule of fields/IEs is defined, reordering costs can be reduced.

Idea of order

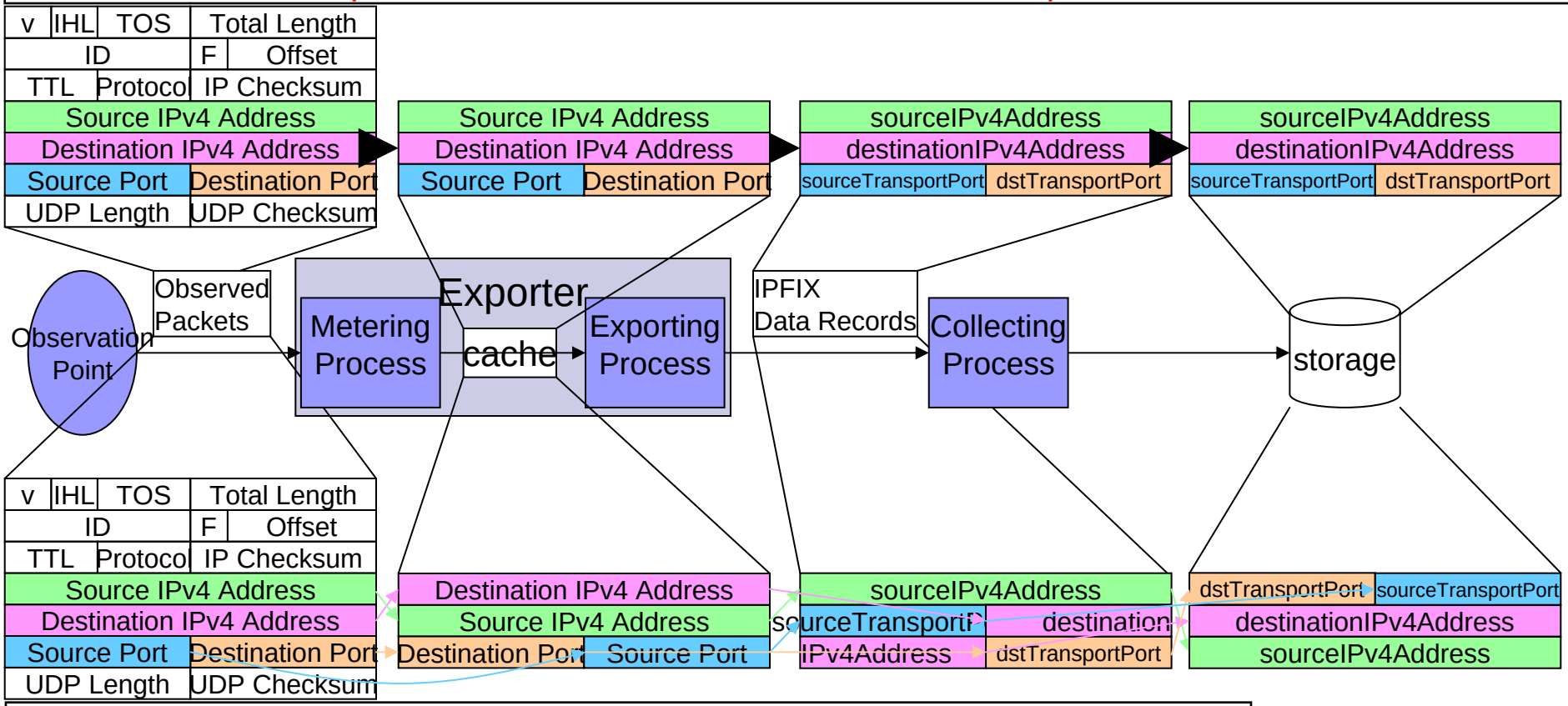
- Idea of order:
  - MPs, EPs and CPs place fields (IEs) in the same order, so it is highly likely that multiple fields will be processed at a time.
    - This reduces reordering costs.
- Order recommended in this presentation
  - Place fields in observed packets in order of protocol header.
  - **Therefore, order of IEs that refer to packets and header fields is recommended.**

	Metering Processes	Exporting Processes	Collecting Processes
Input	Observed packets (network byte order)	Their caches	IPFIX Data Record (network byte order)
Output	(Storing) their caches	IPFIX Data Record (network byte order)	(Storing) files, their DB (real-time analysis)

Example of using same order in MP, EP and CP

Flow Keys: sourceIPv4Address, destinationIPv4Address, sourceTransportPort, destinationTransportPort

Good (ideal) case: Same suggested order,
which refers order of packet header fields used in the cache in Exporter and IPFIX data records



Bad case: Different order used in the cache in Exporter and IPFIX data records

- **If the referential order, which refers to the order of packet fields, is defined, it could, in some cases, lead to increased performance.**
- **If a referential order is undefined, there is no possibility of increased performance.**

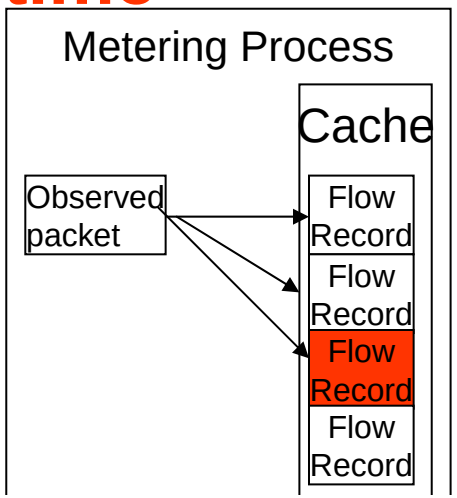
1st idea to improve performance
in environment in which MP, EP, and CP use the same order

Comparison method for multiple fields in Metering Processes (MPs)

NTT Network Service System Laboratories, NTT Corporation

Comparison method for multiple fields in MP (1)

- MP must repeat comparison between existing Flow Records in its cache and new observed packet.
  - To judge whether the new packet belongs to a new flow or an existing one.
- Basically, in this comparison, all fields (IEs) serving as Flow Keys are compared every time.
- **If fields of Flow Records are placed in the same order as packet header fields, MP can compare multiple fields at a time**

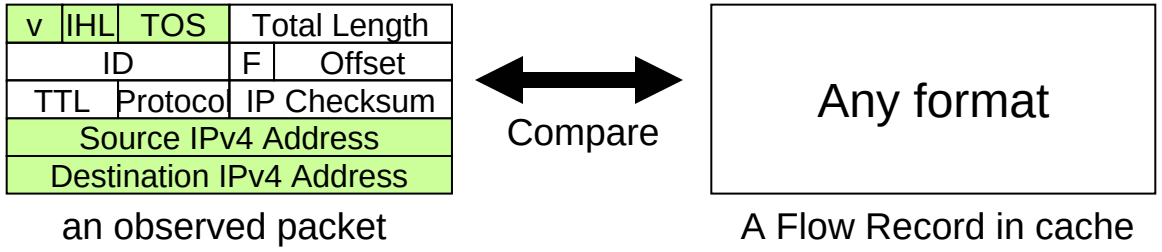


MP repeats comparisons and finds a flow.

Comparison method for multiple fields in MP (2)

Example: Flow Key: Version, IHL, TOS, source Address, destination Address

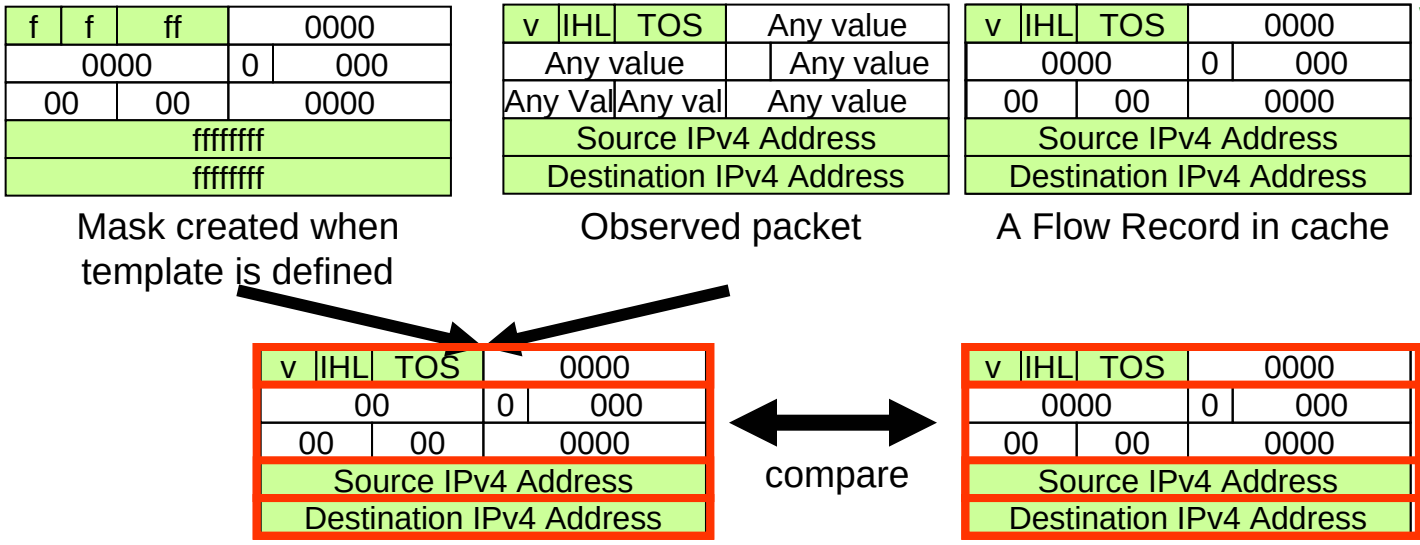
All fields are compared every time (general approach)



When a packet arrives:
5 comparisons
1. ip version
2. IHL
3. TOS
4. Source Address
5. Destination Address

Multiple field comparison (our approach)

Premise: Fields of Flow Records are placed in the referring order as packet header fields



When Template is defined:
Create a Mask

When a packet arrives:
Mask the packet
And
compare these memory areas at the same time
(e.g., memcmp in C language)

Or
1. v + IHL + TOS
2. Source Address
3. Destination Address
(32-bit architecture)

Comparison method for multiple fields in MP (3)

- Number of operations in this method
  - Mask costs smaller than comparison costs.
  - Therefore, this method is effective at increasing performance by reducing the number of comparisons, although it increases mask operations.

	Mask creation	Mask	Comparison
Number of operations	Once in an IPFIX session (when Template is defined)	Depends on the number of observed packets (when packet arrives)	Depends on the number of observed packets and number of flow records in cache

less ⟷ more

- Effective and ineffective cases

v	IHL	TOS	Total Length
ID		F	Offset
TTL	Protoco	IP Checksum	
Source IPv4 Address			
Destination IPv4 Address			

Effective case:
Flow Keys are placed densely

v	IHL	TOS	Total Length
ID		F	Offset
TTL	Protoco	IP Checksum	
Source IPv4 Address			
Destination IPv4 Address			

Ineffective case:
Flow Keys are placed sparsely.

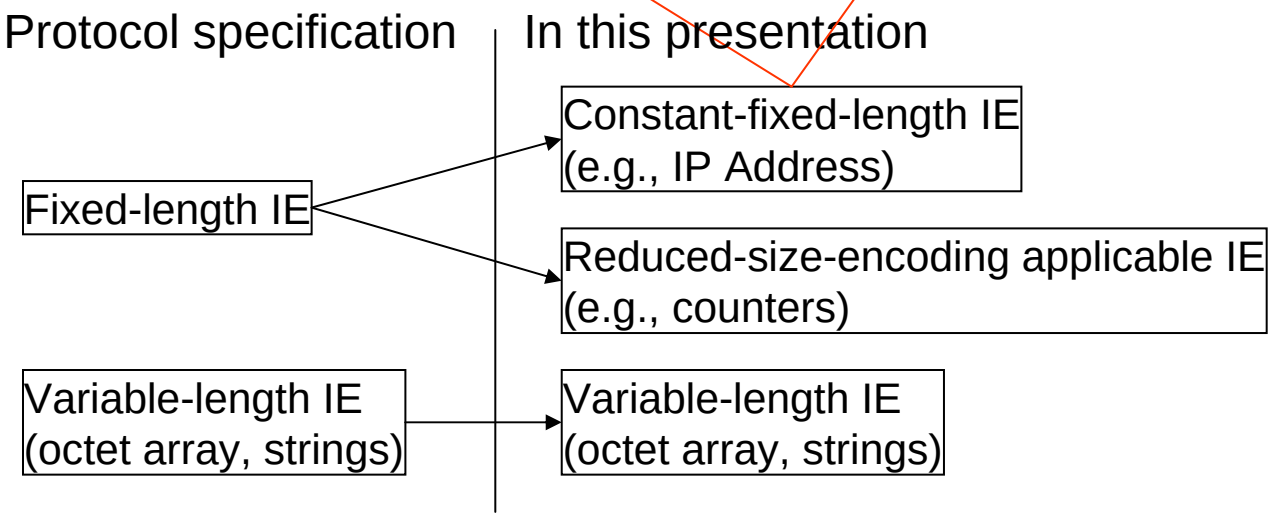
2nd idea to improve performance
in environment in which MP, EP, and CP use the same order

Copy method for multiple fields in Exporting Processes (EPs) and Collecting Processes (CPs)

NTT Network Service System Laboratories, NTT Corporation

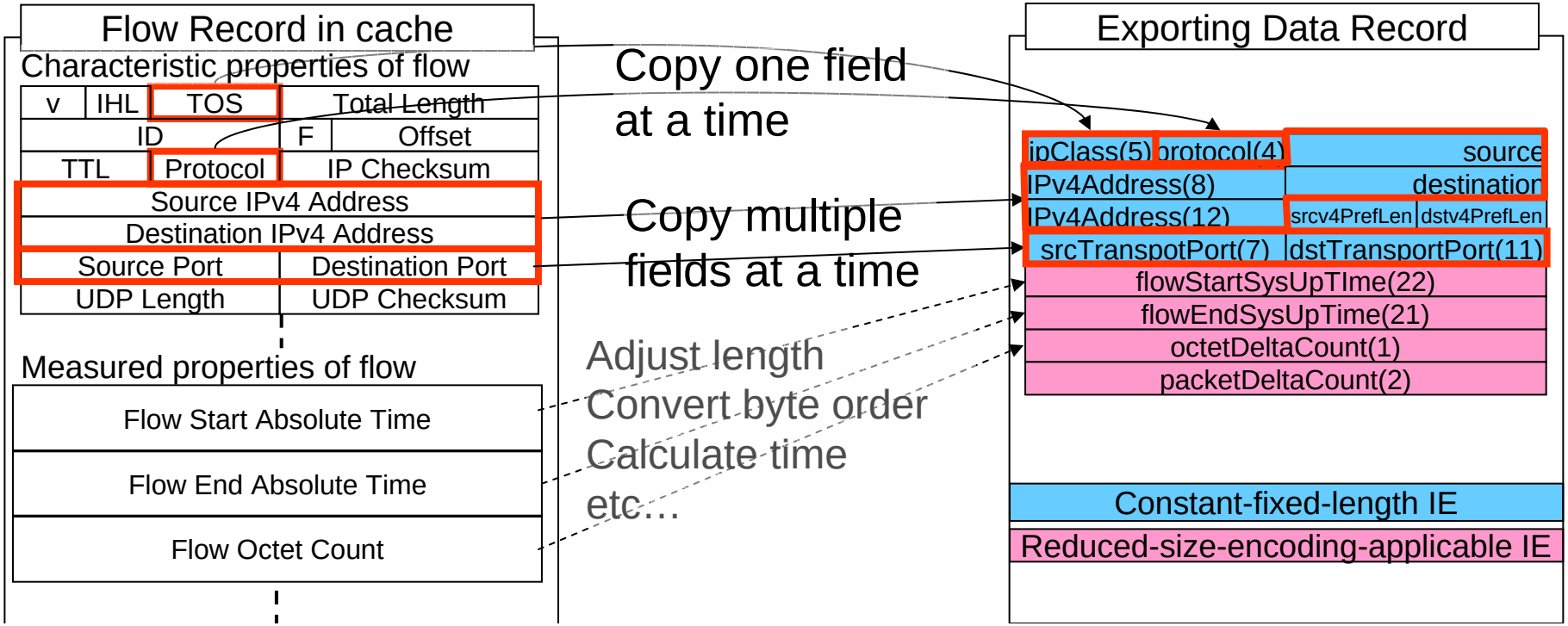
Overview of copy method for multiple fields

- It is a very simple method.
  - If fields in the format of cache and IEs in exporting Data Records are placed in the same order, EPs have a chance to copy multiple adjacent constant-fixed-length IEs at a time.
  - If IEs in received Data Records and fields in Collectors' internal format to store Flow Records are placed in the same order, CPs have a chance to copy multiple adjacent constant fixed-length IEs at a time too.
- IE size classification of IPFIX (terminology in this presentation)



Example of copy method for multiple fields in EP

- Conditions for copying multiple fields
  - Flow Record in cache and Exporting Data Record must use the same order.
  - IEs must have a constant fixed length.
    - Almost all IE characterizing properties of flow are constant fixed length.
  - Byte-orders must be the same.
    - Observed packet and Exporting Data Records use network byte order.
  - IEs for copying multiple fields must be adjacent.



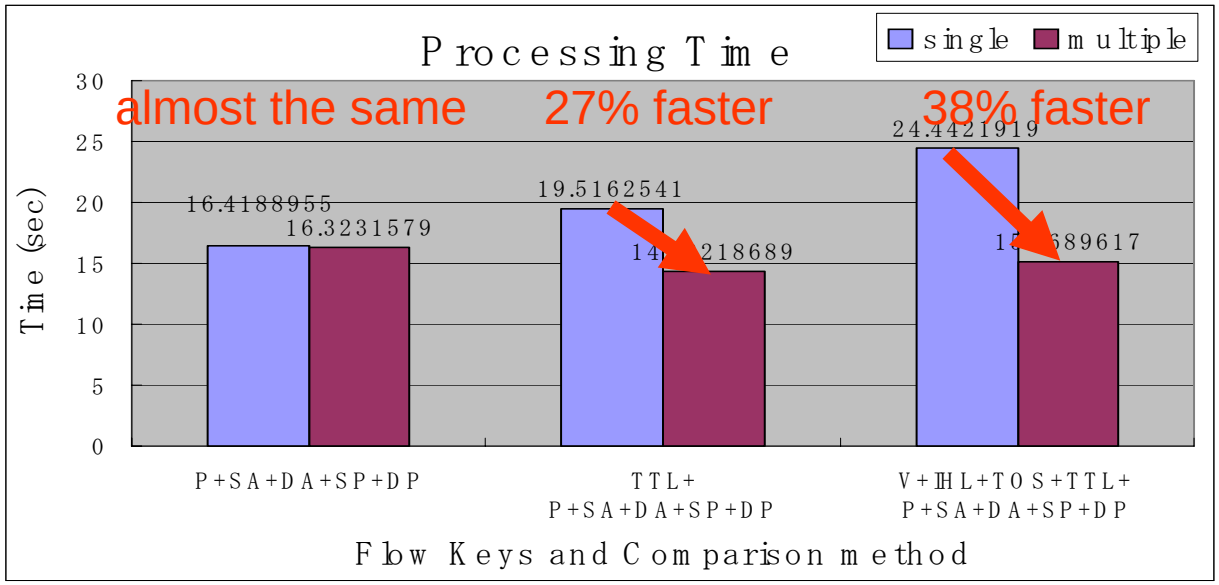
Evaluation & Conclusion

NTT Network Service System Laboratories, NTT Corporation

This material contains an evaluation about only comparison method.

If you want to see an evaluation about copy method, please see a material I talked in past IETF, http://www3.ietf.org/proceedings/07jul/slides/ipfix-10.pdf.

Evaluation of comparison method for multiple fields



v	IHL	TOS	Total Length
ID		F	Offset
TTL	Protoco		IP Checksum
Source IPv4 Address			
Destination IPv4 Address			
Source Port			Dst Port

P+SA+DA+SP+DP

v	IHL	TOS	Total Length
ID		F	Offset
TTL	Protoco		IP Checksum
Source IPv4 Address			
Destination IPv4 Address			
Source Port			Dst Port

TTL+
P+SA+DA+SP+DP

v	IHL	TOS	Total Length
ID		F	Offset
TTL	Protoco		IP Checksum
Source IPv4 Address			
Destination IPv4 Address			
Source Port			Dst Port

V+IHL+TOS+TTL+
P+SA+DA+SP+DP

- **When the density of Flow Key fields is higher, this method works faster.**

Computing environment for the evaluation

- Software Exporter program
  - runs on Intel Xeon 3.06 GHz HT architecture
  - runs on Linux (debian/gnu Linux 4.0)
  - compiled by gcc4
    - optimized option: -O3
- Data used as observed packets:
  - PCAP data published by WIDE project.
  - contains 6,906,333 packets.
  - ftp://mawi.nezu.wide.ad.jp/pub/mawi/samplepoint-B/20060303/200603030100.dump.gz

Conclusion

- Introduced ideas to improve performances of IPFIX processes
  - **Comparison method for multiple fields in MPs**
  - **Copy method for multiple fields in EPs, and CPs**
- These ideas are based on **defining the order rule of IEs/fields**
  - Our Recommendation: IEs/f**ields are placed in the order referring to the packet header fields.**
- The order rule is published as an individual Internet Draft
  - http://tools.ietf.org/id/draft-irino-ipfix-ie-order-03.txt
  - If you agree with these ideas, work with us.